

(19) 대한민국특허청(KR) (12) 공개특허공보(A)

(51) Int. Cl.⁷ (11) 공개번호 특2001-0056748
G06F 17/30 (43) 공개일자 2001년07월04일

(21) 출원번호 10-1999-0058342
(22) 출원일자 1999년12월16일
(71) 출원인 한국전기통신공사
경기 성남시 분당구 정자동 206
(72) 발명자 이운용
경기도의왕시왕곡동세종아파트102-801
이승호
서울특별시서대문구홍제4동인왕산현대아파트106-201
(74) 대리인 이정훈, 이후동

심사청구 : 없음

(54) 어드레스인덱스를 이용한 파일 백업 및 검색 방법

요약

본 발명은 데이터의 저장을 필요로 하는 정보통신서비스 중 데이터의 백업과 백업대상 파일의 축약정보를 추후 검색조건으로 사용하는 경우, 이를 파일 인터페이스를 통해 기술받아 저장함으로써, 추후 조건 검색은 물론 방대한 데이터의 백업내용을 복원시키는 경우 저장된 데이터 어드레스를 이용하여 직접 매체에 접근함으로써 조건 검색은 물론 데이터 백업의 신속한 결과를 얻을 수 있으며, 백업 파일의 확장자를 이용해 파일 복원시 일관성을 확보하기 위한 방법에 관한 것이다.

대표도

도2

명세서

도면의 간단한 설명

도 1은 본 발명에 따른 데이터베이스 정보 예.

도 2는 본 발명에 따른 인터페이스도.

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

본 발명은 어드레스인덱스를 이용한 파일백업 및 검색방법에 관한 것으로, 보다 상세하게는 파일시스템을 사용하는 정보통신서비스 중 파일시스템을 이용한 데이터의 백업과, 백업 작업시 대상 파일의 축약정보를 저장함으로써 추후 백업받은 파일의 조건이나 키를 이용하여 대상화일을 찾거나 복원(Recovery)시키는 경우 물리적 장치의 어드레스를 이용하여 직접 파일에 접근함으로써 신속한 복원처리는 물론 특정 조건에 맞는 파일의 탐색 및 복원시 성능향상을 위한 어드레스인덱스를 이용한 파일 백업 및 검색 방법에 관한 것이다.

일반적으로 정보통신서비스 중 데이터의 전달경로에 따라 정보의 안정적인 저장장치로 데이터베이스시스템을 널리 활용하고 있으나, 데이터의 중요성이 상대적으로 떨어지는 경우에는 파일시스템 자체를 사용한다.

이러한 파일시스템은 데이터의 저장 및 안정적 관리를 위해 별개의 관리기능을 두거나 응용프로그램을 이용하여 데이터를 관리하며 데이터의 관리기능에 반드시 수반되는 기능은 물리적 저장장치의 훼손이나 응용서비스의 장애에 대비한 데이터의 백업 및 복원기능이 포함되며 이러한 기능을 만족시키기 위해 다양한 메커니즘을 이용하여 구축하며, 이러한 백업시스템은 무엇보다도 백업과 복원의 기능중 중요한 요인은 안정적인 메커니즘과 빠른 읽기/쓰기 기능이다.

이러한 속도의 필요성과 안정성 확보방안으로는 단순한 파일시스템을 사용하는 것보다 상용데이터베이스시스템을 사용하는 것이 널리 이용되는데 상용데이터베이스시스템은 수년간 사용되어 왔으므로 시스템의 안전성 측면에서는 새로 개발된 데이터관리시스템보다는 유리하지만 오래되고 널리 쓰인다고 해서 모든 기능이 검증되었다고 단정지을 수 없으며 이는 서비스나 기능의 필요성에 특성화된 데이터베이스시스템

이 많기 때문이다.

기존의 파일시스템 백업은 변경데이터의 데이터 백업이 위주이나 별개의 백업기능이 없다면 전체 파일 백업(즉 디스크나 테이프 또는 저장매체(JukeBox)등)을 하는데, 이때는 백업 대상에 따라 전체를 백업 받는 Full백업을 수행하는 것이 일반적이며, 이전의 백업과 비교하여 변경된 내용만 백업 받는 경우에는 이전의 데이터와 비교할 수 있는 메커니즘이나 식별자가 있어야 한다.

일반적으로 파일시스템의 선택적 백업을 위해서는 변경 데이터를 표시하고 관리하는 기능이 필요하며, 변경여부를 구분할 수 없는 경우는 Full백업을 받아야하고, Full백업데이터의 양이 많아지고 주기가 짧아지면 백업의 속도 및 복원을 위한 검색도구가 제공되는 경우 백업속도도 중요하지만 백업 자료의 위치 및 검색조건에 맞는 데이터의 접근에 처리되는 시간이 오래 걸린다. 상용시스템의 경우에는 Import자료의 일관성을 상용데이터베이스시스템 자체가 유지하지만 오래 시간이 경과된 경우에는 이전상태의 상황 재현이 어려우며 Import기능을 사용하여 데이터를 데이터베이스 시스템내로 복원한 후어야 원하는 검색 조건을 수행할 수 있다.

따라서 파일시스템을 사용하는 경우에는 별도의 백업기능 및 Import기능을 구현해야하며, 데이터의 일관성 및 복원은 상용데이터베이스 시스템을 사용하는 경우보다 일반적으로 어려운데 파일시스템을 모두 다른 파일시스템으로 백업 받을 수는 있지만 이를 하나의 Blob데이터로 취급하여 백업하기 때문에 저장 시 물리적 또는 공간적 어려움이 발생함은 물론 검색시 파일시스템을 복원한 후 이를 대상으로 조건 검색을 하는 등의 문제점이 있었다.

발명이 이루고자하는 기술적 과제

이에 본 발명은 상기와 같은 종래의 문제점을 해결하기 위해 안출된 것으로, 파일시스템은 물론 데이터베이스시스템의 백업 및 복원 시 활용할 수 있는 조건검색 및 메커니즘에 관해 기술하며, 이와 관련된 방법을 제시함으로써 백업시 백업대상의 특정조건 값을 저장함으로써 검색이나 복원시 처리속도를 개선하고 데이터의 일관성을 보장할 수 있도록 한 어드레스인덱스를 이용한 파일 백업 및 검색 방법을 제공함에 그 목적이 있다.

상기와 같은 목적을 달성하기 위한 본 발명은 백업파일정보 생성 및 복원의 일관성확보를 위한 인터페이스의 정보 화일을 구성하는 제 1과정; 상기의 정보에 따라 백업레코드 각각의 검색에 사용될 필드의 정보를 인터페이스 파일로 제공받아 이를 백업저장 대상에 적용할 수 있도록 하는 제 2과정; 상기의 백업 정보에 의한 검색정보 추출 및 물리적 공간으로 저장하고 물리적 위치 및 주소를 검색 조건값과 연계시켜 저장하는 제 3과정; 및 조건값에 의한 검색시 이를 신속하게 처리하기 위한 데이터베이스 구성하는 제 4과정을 구비하는 것을 특징으로 한다.

발명의 구성 및 작용

이하에 첨부된 도면을 참조하여 본 발명의 실시예를 상세히 설명하면 다음과 같다.

먼저 본 발명에 따른 정보검색의 검색 및 조건검색시 빠른 처리를 위한 제약은 다음과 같다.

첫째, 날짜와 관련되는 부문은 모두 숫자로 처리한다. 즉 1월 1일은 1이라는 숫자 값을 갖으며, 12월 31일은 윤년 여부에 따라 365 또는 366의 값을 갖으며, 시작 연도는 데이터베이스로 관리됨으로 특정 연도로 지정하지는 않고 연도 관련 값이 음수이면 백업대상에서 제외하는 방법을 사용하여 Y2K관련문제는 발생하지 않도록 연도는 4자리로 작성한다.

둘째, 백업 및 복원을 기록하는 인터페이스 파일명은 반드시 연도숫자를 인터페이스로 사용하고, 연도가 걸쳐있는(1999.12.12 ~ 2000.1.15) 경우에는 조건값에 따라 동시에 2개의 인터페이스 파일을 조작할 수 있도록 한다.

셋째, 저장정보는 백업 수행 결과 횟수를 나타내며 하루에 발생하는 백업작업은 온라인서비스의 경우에는 1회이며, 배치작업으로 수행하는 경우에는 수회에서 여러회가 될 수 있으며, 이러한 경우 백업작업의 결과가 성공한 횟수를 저장하며 실패한 경우에는 저장하지 않는다.

넷째, 백업정보의 검색을 위해서는 검색조건이 되는 데이터의 애트리뷰트를 지정 받아 저장해야하며, 이때 저장되는 데이터는 반드시 정형화(Formatted)된 자료이어야 한다.

다섯째, 백업작업 수행 시 결과가 실패인 경우에는 일관성확보를 위해 결과파일을 삭제함으로써 일관성을 유지한다.

이러한 조건에 의해 수행한 백업 내용은 파일시스템 백업과 비교하면 기존 파일시스템에서 저장되는 단순 백업과는 달리 저장시 자료의 크기와 저장대상의 축약정보에 해당하는 세부 애트리뷰트의 값을 추출한 후 저장함으로써 추후 검색 시 원본을 대상으로 작업하지 않아도 되는 장점이 있으며, 저장되는 자료의 크기는 해당 백업데이터의 저장위치를 산정하는 의미보다는 다음에 저장될 데이터의 저장 위치를 산정 하는 중요한 위치의 값 의미가 있다.

백업정보를 작성하는데 필요한 도구에는 Unix시스템에서 제공하는 라이브러리함수를 이용하여 대표적으로 Compress방식과 Archive명령어를 사용하며, 백업 대상파일의 각각에 대해서는 Compressing을 수행하여 단위파일을 작성하고(filename.Z), 단위파일을 저장하기 위해서 Archive파일로 다시 작성하는데(filename.a), 이러한 저장메커니즘은 저장되는 데이터의 크기를 줄여줄 뿐 아니라 저장매체의 사용을 효율적으로 줄임으로서, 대용량의 매체를 사용하는 경우 탐색시간의 최소화를 꾀할 수 있는 장점이 있다.

따라서, 추후 백업데이터의 검색을 위해서는 정형화된 자료의 저장이 반드시 필요하며, 검색조건은 날짜나 특정 키 값 또는 조건 값이어야 하며 검색 조건이 단순할수록 저장되는 정보의 양이 줄어들 수 있음으로 연속적으로 조회할 수 있도록 특정단위로 구분하여(메모리 상 작업할 수 있는 임의의 크기) 저장하

는 방법을 사용하는 것이 처리속도를 개선할 수 있고, 이렇게 작성된 백업 데이터의 검색조건 값은 별도의 데이터베이스에 저장할 필요 없이 파일시스템의 데이터베이스로 활용할 수 있다.

보다 안정적으로 백업정보를 저장하는 경우에는 백업 메커니즘(테이프, 디스크, Jukebox 등)의 헤더부문의 일정영역에 검색대상조건을 만족시키는 조건 값을 할당하여 백업 받음으로써 추후 매체장치의 정보를 별도로 표기하지 않더라도 헤더부문을 모아 새로운 데이터베이스의 정보로 구성하도록 만들 수 있는데 검색대상 데이터는 검색 작업 시나 또는 연산 시 메모리 상에 상주시켜 작업함으로써 보다 빠른 결과 값을 얻을 수 있다.

백업 수행에 따른 단위 키 구성은 백업날짜와 백업작업의 결과가 기본적으로 사용되며, 검색조건이 모두 키로 구성될 때를 대비하여 조건 입력 값을 Dynamic Indexing할 수 있도록 구성한다. 이는 검색조건이 문자와 숫자로 입력되는 경우는 ASCII값으로 계산하여 처리할 수 있으나 한글인 경우와 띄어쓰기의 경우에는 Indexing값보다는 이미지 매핑 또는 'Like'항수를 이용하여 처리할 수 있도록 융통성을 제공해야 하기 때문이다.

날짜를 기준으로 하는 연산의 기본방법은 백업작업의 횟수표기 크기와 해당 년도의 일자를 설정한 만큼 곱한 후 데이터파일의 이름을 부여하면 되는데 추후 각 연도를 기준으로 파일 병합을 하는 경우를 대비하여 백업작업 원년을 기준으로 산정 하면 된다.

백업 검색 대상정보를 저장하기 위해서는 저장대상의 데이터의 한 레코드에서 각각의 검색대상 키로 지정된 데이터를 필드 표기에 의한 자릿수만큼을 지정 받아 백업대상의 레코드 수만큼을 저장한다.

백업파일의 형태는 텍스트형태로 저장되며 데이터베이스내의 데이터 포맷은(Date, Fax, Teletex, Time, String, Videotex, Voice 등)변환하거나 파일시스템 형태로 읽어 저장해야 하는데, Unix시스템에서는 변환정보가 ASCII파일로 변환되어 파일의 압축/해제시 무결성을 보장하기 때문이다.

백업정보와 백업매체간의 정보저장 방법은 첫째, 매체저장기의 일정 블록을 나누어 백업축약정보를 저장할 수 있어야 한다. 이때 저장되는 데이터의 크기 산정은 매체 전체의 용량과 백업정보의 단위별크기를 기준으로 계산해야하며 추후 백업정보의 추가적인 쓰기 작업 시 주소를 기준으로 쓰기 동작 위치로 이동하여 해당정보를 저장한다. 이러한 메커니즘은 저장매체의 헤더부문에 저장되어 있는 데이터의 축약 정보를 저장함으로써 추후 전체 백업매체를 대상으로 검색하는 경우 빠른 읽기를 수행할 수 있으며, 기본적으로 시스템에 똑같은 복사 본을 둬으로써 직접적인 백업장치에 접근하지 않고도 조건검색이 이루어질 수 있는 장점을 확보할 수 있다.

백업정보를 이용하여 복원하는 방법에는 단순히 몇몇 조건을 만족시키는 데이터의 추출을 고려할 수 있으나, 전체 데이터의 복원은 데이터베이스시스템을 사용하는 경우에는 테이블 정보가 비어있는 상태로 만든 후 복원하여야 한다. 복원의 일관성확보를 위해서 데이터의 백업 시 저장된 백업 횟수의 숫자만큼을 복원대상의 파일로 간주하고 복원 작업 시 발생되는 오류를 방지하고 일관성을 유지하기 위해 특정 백업파일의 확장자를 인식하여 압축파일 해제 시 백업대상파일 여부에 따라 작업을 수행할 수 있다.

상기와 같은 개념을 바탕으로 구현된 본 발명의 방법이 적용되는 시스템의 개략적인 구성을 살펴보면 도 1에 도시한 바와 같이 백업정보를 저장하고 있는 인터페이스화일과 백업받는 레코드의 검색대상 특정 필드를 지정할 수 있는 인터페이스화일을 사전에 구비한다.

상기와 같이 각각의 인터페이스정보 작성은 인터페이스 파일명이 '1999.idx'라하면 첫 번째부터 3비트를 읽으면 '011'이라는 숫자로 기록되어 있는데 이는 1999년 1월 1일 백업작업을 11번 했다는 의미이며 관련된 Archive화일이 1개이고 그 속에는 Compress한 파일이 11개가 존재한다는 의미이다. 예를들면 1999.0101.001.Z .. 1999.0101.0011.Z 등의 압축화일이 11개라는 의미이다.

이와같은 방법으로 압축된 파일의 내용을 살펴보면 1999.0101.001.Z화일은 압축이전 단계에서 백업정보의 인터페이스 파일에 기술된 정보를 통해 검색대상의 레코드 필드에 대해 저장대상 원본 화일의 일부 필드를 추출하여 정형화된 별도의 데이터베이스 화일을 작성하고, 압축이후 단계에서는 파일의 크기 및 이전에 등록된 물리적 저장의 끝위치(이후 새로 저장될 파일위치)에 저장될 크기를 합산하여 새로운 저장파일의 위치를 지적하고 물리적저장이 완료되면 관련 정보를 파일을 찾기 위한 주소를 해당 키의 특정 필드로 작성하여 저장한다.

이를 바탕으로 추후 검색 및 백업대상에서 추출하는 경우 상기 관련정보의 쓰임새는 검색시 각 필드를 인덱스값으로 작성하여 해쉬방식으로 접근하거나, Linked-List로 작성하여 검색하여 조건값과 일치하는 해당정보의 위치는 물리적이 주소를 기재한 필드를 이용하여 물리적매체 및 위치를 정확하게 지적하여 추출할 수 있다.

물론 검색시 사용한 백업화일의 정보 및 물리적 저장장치의 내용은 HDD나 시스템 자체의 저장장치에 있어야 하며, 장기간 백업정보를 저장하여 관리하는 경우에는 장치 헤더부문에 저장하여 추후 별도의 정보 없이 물리적장치 단독으로 검색 및 저장정보를 활용할 수 있다.

발명의 효과

이상에서 살펴본 바와같이 본 발명에서 백업 저장되는 데이터의 크기 산정은 매체 전체의 용량과 백업정보의 단위별크기를 기준으로 계산해야하며 추후 백업정보의 추가적인 쓰기 작업 시 주소를 기준으로 쓰기 동작 위치로 이동하여 해당정보를 저장한다. 이러한 메커니즘은 저장매체의 헤더부문에 저장되어 있는 데이터의 축약 정보를 저장함으로써 추후 전체 백업매체를 대상으로 검색하는 경우 빠른 읽기를 수행할 수 있으며, 기본적으로 시스템에 똑같은 복사 본을 둬으로써 직접적인 백업장치에 접근하지 않고도 조건검색이 이루어질 수 있는 장점을 확보할 수 있으며, 백업대상 화일의 검색조건을 선별적으로 지정함으로써 빈번하게 사용되는 필드를 별도로 데이터베이스로 구축하는 잇점 외에 백업시스템의 장애 발생시 백업장치 단독으로도 모든 정보를 헤더에서 읽어 처리할 수 있다.

아울러 본 발명의 바람직한 실시 예들은 예시의 목적을 위해 개시된 것으로 당업자라면 본 발명의 사상

과 범위안에서 다양한 수정, 변경, 부가등이 가능할 것이며, 이러한 수정 및 변경 등은 이하의 특허 청구범위에 속하는 것으로 보아야 할 것이다.

(57) 청구의 범위

청구항 1

백업을 수행하는 시스템에 있어서,

백업이전에 필요한 대상 정보의 키 필드를 인터페이스의 정보로 정의하고 관리하는 제 1과정;

상기의 정보에 따라 백업레코드 각각의 검색에 사용될 필드의 정보를 인터페이스 파일로 제공받아 이를 백업저장 대상에 적용하는 제 2과정;

상기의 백업정보에 의한 검색정보 추출 및 물리적 공간으로 저장하고 물리적 위치 및 주소를 검색 조건 값과 연계시켜 저장하는 제 3과정; 및

상기의 조건값에 의한 검색시, 조건검색에 따라 해당 정보를 이용해 저장매체로 직접 접근하여 검색하는 제 4과정을 포함하는 것을 특징으로 하는 어드레스인덱스를 이용한 파일 백업 및 검색 방법.

청구항 2

제 1 항에 있어서,

상기 백업정보는 그 횟수를 기술함으로써 백업 수행에 따른 복원시 대상화일의 갯수를 사전에 계산할 수 있도록 한 것을 특징으로 하는 어드레스인덱스를 이용한 파일 백업 및 검색 방법.

청구항 3

제 1 항에 있어서,

상기 조건검색시 백업정보의 사전 조건은

날짜와 관련되는 부분은 모두 숫자로 표현하고, 백업 및 복원을 기록하는 인터페이스 파일명은 연도숫자를 인터페이스로 저장하고, 백업작업의 결과가 성공한 횟수를 저장하며 실패한 경우에는 저장하지 않고, 검색조건이 되는 데이터의 애트리뷰트를 지정받아 저장하고, 그리고 백업작업 수행시 그 결과가 실패한 경우 결과파일을 삭제하는 것을 특징으로 하는 어드레스인덱스를 이용한 파일 백업 및 검색 방법.

청구항 4

제 1 항에 있어서,

상기 백업정보는 백업 메카니즘의 헤더 부분의 일정영역에 검색대상 조건을 만족시키는 조건 값을 할당하여 저장하는 것을 특징으로 하는 어드레스인덱스를 이용한 파일 백업 및 검색 방법.

도면

도면1

[011] 0230000020013009013009003002 ...
00004400230009200000

파일명 : 1999.dix

파일크기 (3 X 365 by tes)

도면2

A	B	C	D	E	F	G
홍길동	18920123,9201231396588	F,T,I,2,002340				
홍길순	18930212,9302125567378	M,T,I,2,002800				
홍길자					

- A : 성명, A 18 (알파벳 18자리까지)
- B : 생년월일, Data 8 (Data형 8자리만)
- C : 주민등록번호, N13 (Numeric 13자리만)
- D : 성별, A1 (알파벳 1자리)
- E : 테이프 라벨, 디스크 라벨
- F : 저장 섹터
- G : 저장 주소